

# Entropy and enthalpy convergence of hydrophobic solvation beyond the hard-sphere limit

Felix Sedlmeier,<sup>1</sup> Dominik Horinek,<sup>2</sup> and Roland R. Netz<sup>1,a)</sup>

<sup>1</sup>Physik Department, Technische Universität München, 85748 Garching, Germany

<sup>2</sup>Institut für Physikalische und Theoretische Chemie, Universität Regensburg, 93040 Regensburg, Germany

(Received 3 August 2010; accepted 2 December 2010; published online 1 February 2011)

The experimentally well-known convergence of solvation entropies and enthalpies of different small hydrophobic solutes at universal temperatures seems to indicate that hydrophobic solvation is dominated by universal water features and not so much by solute specifics. The reported convergence of the denaturing entropy of a group of different proteins at roughly the same temperature as hydrophobic solutes was consequently argued to indicate that the denaturing entropy of proteins is dominated by the hydrophobic effect and used to estimate the hydrophobic contribution to protein stability. However, this appealing picture was subsequently questioned since the initially claimed universal convergence of denaturing entropies holds only for a small subset of proteins; for a larger data collection no convergence is seen. We report extensive simulation results for the solvation of small spherical solutes in explicit water with varying solute-water potentials. We show that convergence of solvation properties for solutes of different radii exists but that the convergence temperatures depend sensitively on solute-water potential features such as stiffness of the repulsive part and attraction strength, not so much on the attraction range. Accordingly, convergence of solvation properties is only expected for solutes of a homologous series that differ in the number of one species of subunits (which attests to the additivity of solvation properties) or solutes that are characterized by similar solute–water interaction potentials. In contrast, for peptides that arguably consist of multiple groups with widely disperse interactions with water, it means that thermodynamic convergence at a universal temperature cannot be expected, in general, in agreement with experimental results. © 2011 American Institute of Physics. [doi:10.1063/1.3530585]

## I. INTRODUCTION

The question of protein folding and stability is one of the fundamental problems in biology.<sup>1</sup> Despite intense scientific efforts since Anfinsen's pioneering work,<sup>2</sup> it is not yet completely clear what the dominant forces are that govern protein stability.<sup>3,4</sup> The hydrophobic effect has long been suspected to play a major role.<sup>5</sup> It was therefore seen as a major breakthrough, when the observation of convergence of the denaturation entropy per residue of a whole class of proteins at a universal temperature allowed to draw a close analogy with the solvation properties of simple hydrophobic compounds.<sup>6,7</sup>

It was first observed by Privalov and Khechinashvili<sup>8,9</sup> that the entropies and enthalpies of denaturation per residue for a number of proteins, if extrapolated to high temperatures, converge to a common value at a temperature near  $T_S^* \approx T_H^* \approx 110^\circ\text{C}$ . Baldwin<sup>10</sup> noticed that at roughly the same temperature the entropy of dissolution of several liquid hydrocarbons<sup>11</sup> extrapolates to zero (while the convergence temperature of the solvation enthalpy is much lower). A similar convergence has been observed<sup>6,7</sup> for the dissolution of saturated hydrocarbon gases,<sup>12–16</sup> noble gases,<sup>12,17,18</sup> cyclic dipeptides,<sup>19,20</sup> and gaseous alcohols.<sup>21</sup> These observations

stirred an intense discussion about the interpretation and implications of these findings for the role of the hydrophobic effect in protein folding,<sup>7,19,21–25</sup> which we will summarize briefly further below.

Theoretically, the temperature dependence of hydrophobic hydration has been studied by various methods,<sup>26–40</sup> mostly focusing on the solvation of simple model solutes such as hard spheres<sup>27,28,30–33,35,39</sup> or Lennard-Jones particles.<sup>26,29,34,37,40</sup> Using a Gaussian model for particle number fluctuations in atomic-scale cavities in conjunction with simulated water radial distribution functions, convergence temperatures of about  $T_S^* \approx 400$  K for small hard-sphere solutes of different radii, close to the experimental results, were obtained.<sup>27,28,31</sup> The model allowed a simple interpretation of this phenomenon in terms of particle number fluctuations, water density variation, and to a limited extent the isothermal compressibility of water.<sup>27,28</sup> Results from the revised scaled-particle theory<sup>39</sup> show that the convergence temperature  $T_S^*$  for larger sphere radii starts to depend on the size of the solutes, but for typical solute radii around 3 Å convergence temperatures in the range  $\approx 370$ –400 K are observed.

Despite this seemingly good agreement of the entropy convergence temperatures obtained from theoretical models for hard-sphere solvation with the experimental results for both proteins and simple hydrophobic compounds, a few open questions remain:

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: netz@ph.tum.de.

- (i) For simple hydrophobic compounds, the convergence temperatures of solvation entropy and enthalpy widely differ, the latter being lower by typically  $100^\circ$ , while the putative entropy and enthalpy convergence temperatures for proteins are the same.<sup>7,10</sup> To account for this, it is assumed that the solvation of hydrophilic groups contributes to the enthalpy but not to the entropy so that the solvation entropies of proteins and hydrophobic compounds are similar but not the enthalpies.<sup>22</sup> In contradiction to this assumption and not surprisingly, we find a continuous change of the solvation entropy as one goes from hydrophobic to hydrophilic solutes and in all cases the solvation entropy does not strictly vanish.
- (ii) The theoretically predicted convergence of hard-sphere solvation entropies at about 400 K was obtained based on simulation results for simple point charge (SPC) water simulated at the experimental liquid–vapor coexistence density.<sup>27</sup> As we show in this paper, at ambient pressure of 1 bar, which seems the relevant condition when comparing to protein data, the entropy convergence temperature for hard-sphere solvation goes down to 340 K, in disagreement with both protein and simple hydrocarbon data.
- (iii) Perhaps most strikingly, Robertson and Murphy<sup>41</sup> showed that for a larger set of proteins neither entropy nor enthalpy convergence is observed (see Fig. 1, circles), although the polar and nonpolar surface areas per residue, which are buried upon folding, show the same trends as for the smaller set of proteins originally studied (denoted by triangles).<sup>7</sup>

In this paper we do not directly consider protein denaturation thermodynamics, but rather investigate the solvation of a more general class of spherical solutes. It has been shown that attractive solute–water interactions have a pronounced influence on the solvation thermodynamics.<sup>29,31,33,34,38</sup> We therefore systematically study the dependence of the entropy and enthalpy convergence temperature for spherical solutes on the solute–water interaction potential. We use a Buckingham double exponential potential with four free parameters to vary the stiffness of the repulsive part, the range and strength of the attractive part of the interaction, as well as the solute radius independently. Solvation free energies are obtained from molecular dynamics (MD) simulations by the particle insertion method. We mainly use the popular extended simple point charge (SPC/E) water model, but also compare our results with the SPC and the TIP5P parameters to check the influence of the water model<sup>37,40</sup> on the solvation properties. As main results, we show that thermodynamic convergence exists, but in a much more restricted sense than previously believed:

- (i) For hard-sphere solutes, convergence for different sphere radii in the range of 0.28–0.35 nm (matching the previously studied radii as relevant for noble gases and methane)<sup>27</sup> exists but at a fixed pressure of 1 bar occurs for the entropy at roughly 350 K and for the enthalpy at an extrapolated temperature of about 205 K. Experimentally, the entropy convergence temperature for no-

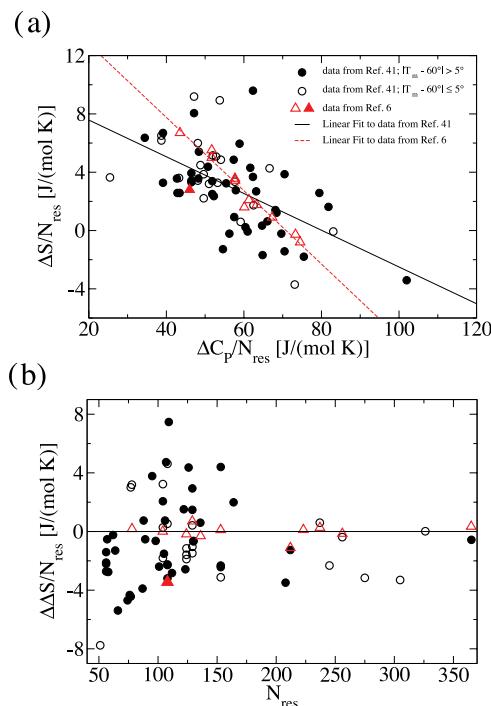


FIG. 1. (a) Entropy change  $\Delta S$  plotted against the change in isobaric heat capacity  $\Delta C_p$  upon denaturation per residue extrapolated to  $25^\circ\text{C}$  for different sets of proteins. The small set (triangles) is taken from Ref. 6. Open triangles denote the proteins, that have been used in the analysis of Ref. 7 to argue for the occurrence of entropy convergence in protein denaturation. The data point for parvalbumin (filled triangle) was included in the data set of Ref. 6 (see Table I of Ref. 6) but not shown in the graph and not considered in the data fitting of Ref. 7. The data for the larger set of proteins (circles) are taken from Ref. 41. This set is separated into proteins with denaturation temperatures  $T_m$  in the interval  $55^\circ\text{C} < T_m < 65^\circ\text{C}$  (open circles) and  $T_m$  outside this interval (filled circles). The black line is a linear fit to the data set of Ref. 41, the dashed line to the open triangles. (b) Deviation  $\Delta\Delta S$  of the unfolding entropy per residue from the linear fit for the small set [dashed line in (a)] as a function of the number of residues  $N_{\text{res}}$  of the proteins. No systematic trend as a function of the protein size is seen. Also, there is no qualitative difference between the set of proteins with a denaturation temperature close to  $60^\circ\text{C}$  (open circles) and far away from  $60^\circ\text{C}$  (filled circles).

ble and saturated hydrocarbon gases is located around 385 K. At densities corresponding to the experimental liquid–vapor coexistence of water, the entropy convergence temperature is raised to about 400 K and thus becomes closer to the experimental entropy convergence temperature of gas solvation and protein unfolding, in agreement with previous theoretical findings.<sup>27</sup>

- (ii) Making the repulsive potential softer and thus more realistic, further decreases the convergence temperatures (obtained for different radii) and thus makes the comparison with experiments even worse.
- (iii) On the other hand, including an attractive component to the solute–water interaction increases the convergence temperature (for different solute radii) up to experimentally relevant values.
- (iv) Since in protein denaturing, chemical groups that differ in radius and stiffness of the repulsive potential part as well as in attraction strength are solvated, we conclude that convergence of solvation entropies and enthalpies is not expected, in general, as was suggested before<sup>39</sup>

and in agreement with the above mentioned more recent experimental results.<sup>41</sup>

- (v) For series of simple model compounds, on the other hand, convergence is expected for certain cases: if members of the series differ in the number of one constituent, convergence is obtained and attests to the additivity of solvation properties. Likewise, if the properties of the solute-water potential (such as repulsive stiffness, attraction strength) are similar but the radius differs (as is the case to a certain extent for noble gases) convergence is also expected.

As a word of caution, we note that convergence properties depend sensitively on the water parameters used, which simply reflects the fact that one is dealing with higher-order derivatives of the free energy. Thus, they are very sensitive to small force-field variations and imperfections. The actual numbers quoted by us should thus be considered with care. Likewise, we only study small radii, for which more complicated effects associated with the crossover to the solvation of planar interfaces is not yet reached.<sup>42</sup>

In Sec. II, we review basic notions of thermodynamic convergence of solvation and protein denaturation, in Sec. III we describe in detail the modeling of the solutes and the simulation and analysis methods employed. In Sec. IV, we present the results and discuss the dependence of the entropy and enthalpy convergence on the solute–water interaction details and the water model employed. In Sec. V we summarize and conclude.

## II. THERMODYNAMICS OF CONVERGENCE

Assuming a temperature independent heat capacity change  $\Delta C_p$  upon solvation of a solute or denaturation of a protein, from  $T \partial \Delta S / \partial T = \Delta C_p$  and  $\partial \Delta H / \partial T = \Delta C_p$ , the entropy and enthalpy changes follow as<sup>43</sup>

$$\Delta S(T) = \Delta S(T_0) + \Delta C_p \ln \left( \frac{T}{T_0} \right), \quad (1)$$

and

$$\Delta H(T) = \Delta H(T_0) + \Delta C_p (T - T_0), \quad (2)$$

where  $T_0$  is an arbitrary reference temperature. We will show that this assumption is approximately true for the systems studied by us. If entropy convergence occurs at a temperature denoted by  $T_S^*$  for a group of species, then a plot of  $\Delta S(T)$  versus  $\Delta C_p$  at a fixed temperature  $T$  yields a straight line, whose slope  $\ln(T/T_S^*)$  is determined by the convergence temperature  $T_S^*$  and whose intercept  $\Delta S(T_S^*)$  gives the entropy change at the convergence temperature.<sup>7</sup> Similarly, a plot of  $\Delta H$  versus  $\Delta C_p$  yields a straight line with the slope  $T - T_H^*$  and intercept  $\Delta H^*$ , if enthalpy convergence is observed. Note that such a plot allows to extract the convergence temperature and the convergence entropy or enthalpy based on the assumption of a constant heat capacity in a straightforward manner (but it does not allow to check whether the heat capacity is actually constant). In Fig. 1 we show such a plot for the original set of proteins<sup>6</sup> used in the analysis of Murphy *et al.*<sup>7</sup> (open triangles) including a linear fit to the data. The data in-

TABLE I. Convergence temperatures  $T_S^*$  and  $T_H^*$ , entropies  $\Delta S^*$  and enthalpies  $\Delta H^*$  obtained from the plots shown in Fig. 5 for different reference temperatures  $T$ .

T [K]	$T_S^*$ [K]	$\Delta S^*$ [J/(mol K)]	$T_H^*$ [K]	$\Delta H^*$ [kJ/mol]
300	347.2	−17.8	205.0	−3.1
320	347.4	−17.1	204.7	−3.8
340	347.5	−16.8	204.5	−4.5
360	347.5	−16.8	204.2	−5.4
380	347.5	−17.1	204.0	−6.4

indicated by the filled triangle corresponding to the denaturing of the protein parvalbumin were listed in Ref. 6 (see Table I of Ref. 6) but neither included in the data fitting nor in the graph of Ref. 7. Note that the data in this plot have been extrapolated from the denaturing temperature by assuming the relations Eqs. (1) and (2) to hold, an assumption which for proteins is, in general, difficult to validate since the denaturing temperature can only be varied by adding denaturant or pH change (and thus possibly causing side effects) in a restricted range.

Murphy and Gill<sup>19,23</sup> and Lee<sup>22</sup> showed that convergence of some thermodynamic observable, e.g., the entropy of denaturation of a set of proteins or the entropy of dissolution  $\Delta S$  within a homologous series of compounds, can generally be expected if that observable is linearly dependent on the variable  $X$  (e.g., the number of hydrophobic groups in a homologous series of hydrocarbons) which distinguishes the elements of the series,

$$\Delta S(T) = \Delta S_R(T) + X \Delta S_X(T), \quad (3)$$

where  $\Delta S_X(T)$  is the entropy change associated with the variable  $X$  and  $\Delta S_R$  is the residual entropy change due to the molecular parts that are common within the series. From the relation  $C_p = T \partial S / \partial T$  we obtain a linear dependence of the heat capacity change on  $X$ ,

$$\Delta C_p = \Delta C_R + X \Delta C_X, \quad (4)$$

where we assume all heat capacity contributions to be independent of temperature. Then, by eliminating  $X$  from Eqs. (3) and (4), we obtain

$$\Delta S(T) = \Delta S_R(T) + (\Delta C_p - \Delta C_R) \frac{\Delta S_X(T)}{\Delta C_X}, \quad (5)$$

i.e., a linear dependence of  $\Delta S(T)$  on  $\Delta C_p$ . A sufficient condition for both Eqs. (3) and (4) is the linear dependence of the free energy change  $\Delta F$  on  $X$ .<sup>22</sup> The argument presented above is equivalent to the earlier observation of Sturtevant<sup>44</sup> that the ratio of entropy change and heat capacity change upon dissolution of several hydrophobic compounds is constant. Comparison with Eq. (1) shows that the convergence temperature, given by the slope,

$$\ln \left( \frac{T}{T_S^*} \right) = \frac{\Delta S_X(T)}{\Delta C_X}, \quad (6)$$

depends solely on quantities associated with the variable  $X$ . Equation (5) further implies that the entropy at convergence is equal to the residual entropy,  $\Delta S(T_S^*) = \Delta S_R(T_S^*)$ . Therein

lies the importance of convergence phenomena, since they enable one to separate different contributions to the observable under study.

In the case of protein denaturation it is not *a priori* clear what the distinguishing variable  $X$  is and therefore how the observed convergence should be interpreted.<sup>7,19,21–23</sup> Murphy *et al.*<sup>7,23,45</sup> argue that for globular proteins the polar surface area per residue that is exposed to water upon unfolding is constant, while the nonpolar surface area per residue varies and thus is taken proportional to  $X$ .<sup>45,46</sup> This implies that the contribution of the nonpolar protein surface vanishes at  $T_S^*$ , while the residual entropy is due to solvation of polar groups and configurational entropy, and corroborates the analogy with the dissolution of other hydrophobic compounds.<sup>7,10</sup> Accordingly, the similar convergence temperature for the different compounds is taken as an universal feature of hydrophobic hydration.<sup>7,21,23</sup> Lee<sup>22</sup> originally suggested a slightly different assignment, normalizing the unfolding entropy by the buried surface area instead of the number of residues and taking the fraction of nonpolar surface area exposed upon unfolding as the variable  $X$ . In that case the convergence is observed at the temperature, at which the polar and nonpolar contributions to the entropy are equal. Both approaches essentially can be viewed as different definitions of the hydrophobic contribution to denaturation.<sup>45</sup>

One notes that the mechanism for entropy convergence discussed in Eqs. (3)–(6) obviously rests on the assumption of additivity of the contributions of individual chemical groups to the total solvation entropy and enthalpy, an assumption which has been independently verified.<sup>13,47,48</sup> The other, more serious, assumption is that it is sufficient to sort the solvation contributions of all chemical groups into just two groups, hydrophilic and hydrophobic, characterized by the variable  $X$ , an assumption which we show not to be entirely true in this paper. In fact, the concept of entropy convergence treated theoretically by the information theory approach and revised scaled particle theory in some sense goes beyond the additivity concept outlined above,<sup>19,22,23</sup> since here it is shown that the entropy of solvation of hard spheres converges to a universal value at one temperature for a whole range of different sphere radii. The importance of this observation lies in the fact that the hydrophobic groups that are solvated when a protein denatures in general have different radii. However, the convergence entropy is nonzero for hard-sphere solvation so that according to the assignment of Lee<sup>22</sup> one will in general not find a unique temperature where the hydrophobic contributions to the solvation entropy cancel the polar contributions for different proteins. Even more importantly, individual amino acids differ in more properties than just their radii, since details of the solute-water potential are expected to vary between different species, which adds additional variation to the expected convergence temperatures. Again, all these reservations serve to explain that entropy convergence for protein denaturation is not really expected, in accord with actual experimental observations (see the complete set of protein data in Fig. 1).<sup>41</sup>

The convergence of solvation entropies and enthalpies of liquid hydrocarbons (e.g., alkanes) simply reflects the additivity of solvation of individual linked methyl groups and thus is

not really related to the main point of this paper. An additional complication comes in because transfer studies from liquid hydrocarbons to aqueous solution (as well as protein denaturation studies) involve the liquid (or protein core) reference state, which is difficult to treat theoretically. Such complications are obviously absent for solvation studies with noble or hydrocarbon gases, which are more closely related to our calculations. In this respect it is important to realize that the entropy convergence of noble gases is far from perfect,<sup>39</sup> a fact that is very much in line with our findings.

### III. METHODS AND MODELING

#### A. Water models

For all molecular simulations including water, the choice of the water model is crucial since all results will quantitatively depend on that choice. One of the most common water models is the SPC/E model,<sup>49</sup> which is used for most of the simulations in the current work. For an estimate of the robustness of our results, we also perform simulations using the SPC (Ref. 50) and the TIP5P (Ref. 51) water models. SPC and SPC/E are both rigid three-site models, differing only in the magnitude of the partial charges positioned on the hydrogen and oxygen atoms. In the TIP5P model, whose geometry is also rigid, the negative charge is placed on two additional interaction sites positioned on the edges of a slightly distorted tetrahedron formed together with the hydrogen atoms. In all three models dispersion interactions are accounted for by a single Lennard-Jones site centered on the oxygen atom.

#### B. Solute-water potential

To study the influence of the potential stiffness and the role of attractive interactions on the solvation properties of spherical solutes we choose a Buckingham double exponential potential of the form

$$U_{S-OW}(r) = A e^{-B(r-\sigma)} - C e^{-D(r-\sigma)}, \quad (7)$$

where  $r$  is the distance between solute center and water oxygen,  $\sigma$  is the size of the solute,  $B$  and  $D$  are the inverse decay lengths of the repulsive and attractive parts of the potential, respectively, and  $A$  and  $C$  are constants chosen as follows: in the case of a purely repulsive solute  $C = 0$  and  $A = k_B T_0$  with  $T_0 = 300$  K. This defines the radius of the particle as the separation, where the interaction potential equals  $k_B T_0$ , that is,  $U_{S-OW}(\sigma) = k_B T_0$ . The decay length of the repulsive part is varied in the range  $1/B = 0.005$ – $0.03$  nm with an increment of  $0.001$  nm. The upper limit  $0.03$  nm corresponds to the decay length used in the parameterization of a Buckingham exponential-6 potential of a water oxygen atom<sup>52</sup> and should therefore give a realistic estimate for the stiffness of an interatomic potential. In the limit of vanishing decay length,  $1/B \rightarrow 0$ , one obtains a hard-core potential,

$$U_{S-OW}^{HC}(r) = \begin{cases} \infty, & \text{for } r < \sigma \\ 0, & \text{for } r > \sigma \end{cases}, \quad (8)$$

which we also study as a limiting case.



For the case of an attractive solute the constants  $A$  and  $C$  are chosen such that (i)  $U_{S-OW}(\sigma) = k_B T_0$  and (ii)  $\min(U_{S-OW}(r)) = -\epsilon_0$ , where  $\epsilon_0$  is the depth of the potential. We vary both the range of the attractive part,  $1/D = 0.05, 0.1$ , and  $0.2$  nm, and the depth of the potential,  $\epsilon_0 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ , and  $1.3$  kJ/mol. An illustration of the shape of the potential for various different parameters is given in Fig. 2. The advantage of such a double exponential potential over the more commonly used Lennard-Jones 12-6 potential is that the solute size  $\sigma$ , the decay lengths of the repulsive and attractive parts  $1/B$  and  $1/D$ , and the potential depth  $\epsilon_0$  can be varied independently, whereas in a Lennard-Jones potential there are only two free parameters. We note that the radius  $\sigma$  of the solute is an exclusion radius, that is, it is the radius around the center of the solute from which the center of the water oxygen atom is effectively excluded by the repulsion. We neglect all interactions between the solute and the water hydrogens.

### C. Test particle insertion

The key quantity characterizing the solvation of a solute is its solvation free energy, that is, the free energy necessary

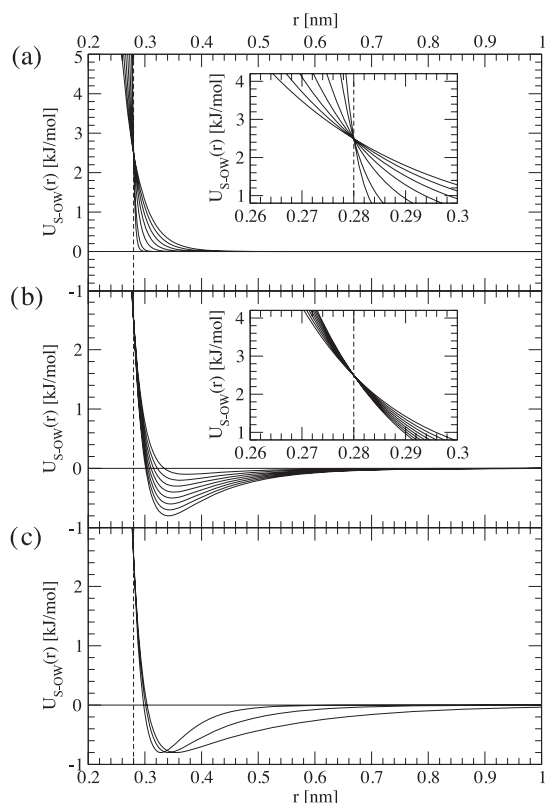


FIG. 2. Illustration of the solute–water interaction potential  $U_{S-OW}(r)$  [see Eq. (7)] for a spherical solute of radius  $\sigma = 0.28$  nm and various potential shapes. (a) Purely repulsive potential with varying decay lengths  $1/B = 0.005, 0.01, 0.015, 0.02, 0.025$ , and  $0.03$  nm. (b) Attractive potentials with  $1/B = 0.02$  nm, attraction decay length  $1/D = 0.05$  nm, and varying potential depth  $\epsilon_0 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ , and  $0.8$  kJ/mol. (c) Attractive potentials with  $1/B = 0.02$  nm,  $\epsilon_0 = 0.8$  kJ/mol, and varying range  $1/D = 0.05, 0.1$ , and  $0.2$  nm. Vertical lines indicate the radius  $\sigma$  of the solute, defined by  $U_{S-OW}(\sigma) = k_B T_0$  with  $T_0 = 300$  K. The insets in (a) and (b) show enlarged views of the potential.

to bring the solute from the pure ideal gas phase into aqueous solution. The solvation free energy per solute is equal to the solutes excess chemical potential  $\mu_{ex}$  in the solution. From the temperature derivative of the excess chemical potential one obtains the solvation entropy  $\Delta S$  per solute,

$$\Delta S(T) = -\frac{\partial \mu_{ex}(T)}{\partial T}. \quad (9)$$

We determine the excess chemical potential of the solutes under study by the Widom particle insertion method.<sup>53</sup> In the case of an isobaric–isothermal ensemble it is given by<sup>54</sup>

$$\mu_{ex} = -k_B T \ln \left( \frac{\langle V e^{-\beta \Delta U} \rangle}{\langle V \rangle} \right), \quad (10)$$

where  $\Delta U$  is the potential energy of the interaction between the solute and the solvent,  $V$  is the volume of the system,  $\beta = 1/(k_B T)$ , and the angle brackets denote an isobaric–isothermal average over configurations of the system without the solute.

### D. Isothermal compressibility

The isothermal compressibility of the SPC/E water model is determined by a finite difference method,<sup>55</sup>

$$\kappa_T = \frac{1}{\rho} \left( \frac{\partial \rho}{\partial p} \right)_T \approx \frac{\Delta \ln(\rho)}{\Delta p} = \frac{\ln(\rho_2/\rho_1)}{p_2 - p_1}. \quad (11)$$

To evaluate Eq. (11) the system is simulated in an NVT (constant particle number, volume, and temperature) ensemble with the densities  $\rho_{1,2} = \rho \pm 0.04$  kg/l and the resulting pressures  $p_{1,2}$  are sampled. Here  $\rho$  is the equilibrium density at a pressure of  $p = 1$  bar, as obtained from previous MD simulations.

### E. MD simulations

The trajectories for the insertions are generated by molecular dynamics simulations using the GROMACS simulation package.<sup>56,57</sup> Systems consisting of 895 water molecules in the case of the SPC/E and SPC water models and 878 water molecules in the case of the TIP5P water model are simulated in a cubic box with periodic boundary conditions. At  $T = 300$  K this corresponds to a box size of roughly  $3.0 \times 3.0 \times 3.0$  nm<sup>3</sup>. We perform simulations at temperatures of  $T = 280, 300, 320, 340, 360, 380, 400, 420, 440, 460, 480$ , and  $500$  K and at a pressure of  $p = 1$  bar. Some additional simulations are performed at the experimental liquid–vapor coexistence density for SPC water for a system containing 256 molecules. At each temperature each system is equilibrated first in an NVT ensemble for  $t = 50$  ps and then in an NPT (constant particle number, pressure, and temperature) ensemble for  $t = 100$  ps. Production runs are performed subsequently for  $t = 10$  ns and configurations are saved each ps. To calculate the isothermal compressibility, additional simulations in an NVT ensemble with the densities  $\rho_{1,2} = \rho \pm 0.04$  kg/l, where  $\rho$  is the average density of the previous NPT simulations, are done for  $t = 10$  ns. A Berendsen weak coupling thermostat and barostat<sup>58</sup> with a relaxation time of  $\tau = 1.0$  ps is used for temperature and pressure control. All

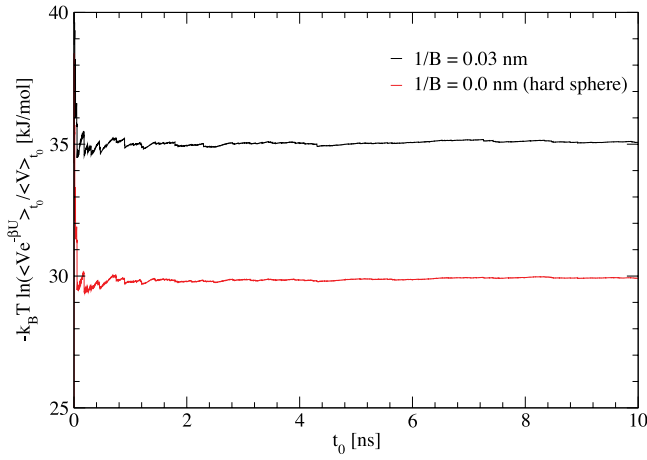


FIG. 3. Running average of the excess chemical potential calculated by the particle insertion method according to Eq. (10) for a purely repulsive solute of radius  $\sigma = 0.345$  nm with a potential stiffness of  $1/B = 0.03$  nm (black curve) and for a hard-core solute of the same size (red line), where the average  $\langle \cdot \rangle_{t_0}$  is taken over all saved configurations with  $t < t_0$ . The insertions are done in a box containing 895 SPC/E water molecules simulated at  $T = 300$  K. In each configuration of the trajectory  $n_{\text{ins}} = 10^7$  insertions are performed for the finite potential stiffness and  $n_{\text{ins}} = 2 \times 10^6$  for the hard-core solutes. It is seen that the excess chemical potential is well converged after 10 ns.

nonbonded interactions are cut off at a radius of  $r_c = 0.9$  nm. Long-range electrostatic interactions are treated by the particle mesh Ewald method<sup>59,60</sup> with tinfoil boundary conditions. An analytical cutoff correction for the Lennard-Jones interaction is applied to both energy and pressure.<sup>61</sup> For each saved configuration  $n_{\text{ins}} = 10^7$  insertions are performed in the case of a finite potential stiffness and  $n_{\text{ins}} = 2 \times 10^6$  in the case of the hard-sphere solutes. Figure 3 shows the running average, that is, an average including only configurations up to a certain time  $t_0$ , of the excess potential [Eq. (10)] for a purely repulsive solute of radius  $\sigma = 0.345$  nm with a potential stiffness of  $1/B = 0.03$  nm (black curve) and a hard-core solute of the same size (red curve) in SPC/E water at  $T = 300$  K. It is seen that the excess potential is well converged after 10 ns.

## IV. RESULTS AND DISCUSSION

### A. Results for hard-core solutes

In this paper we determine the solvation free energy  $\mu_{\text{ex}}$  as a function of the temperature for solutes with four different radii  $\sigma = 0.28, 0.31, 0.33$ , and  $0.345$  nm, as used by Garde *et al.*<sup>27</sup> to model solvation of neon, argon, methane, and xenon, and various potential shapes [see Eq. (7) and the following text]. Figure 4(a) shows the solvation free energies obtained by particle insertion (symbols) for hard-core solutes in SPC/E water. In qualitative agreement with previous results,<sup>27</sup>  $\mu_{\text{ex}}$  has a maximum around 380 K. We fit the data by the form,

$$\mu_{\text{ex}}(T) = a + bT - cT^2 - dT \ln(T), \quad (12)$$

leading to the solvation entropy,  $\Delta S = -\partial \mu_{\text{ex}} / \partial T$ ,

$$\Delta S(T) = d - b + 2cT + d \ln(T), \quad (13)$$

the solvation enthalpy,  $\Delta H = \mu_{\text{ex}} + T \Delta S$ ,

$$\Delta H(T) = a + dT + cT^2, \quad (14)$$

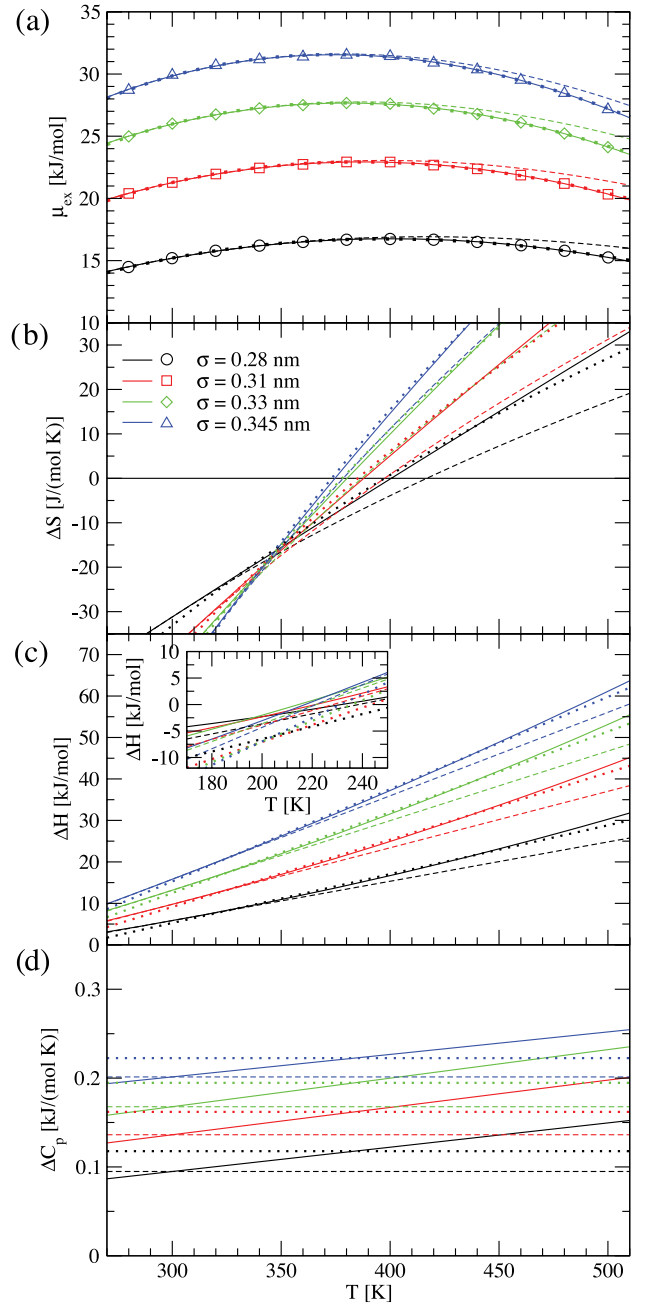


FIG. 4. (a) Excess chemical potential  $\mu_{\text{ex}}$ , (b) entropy  $\Delta S$ , (c) enthalpy  $\Delta H$ , and (d) isobaric heat capacity  $\Delta C_p$  of differently sized hard-core solutes in SPC/E water as a function of the temperature. The solutes have radii of 0.28 nm (circles, black lines), 0.31 nm (squares, red lines), 0.33 nm (diamonds, green lines), and 0.345 nm (triangles, blue lines), corresponding roughly to the size of neon, argon, methane, and xenon. Excess chemical potentials are determined by test particle insertion (symbols). Solid lines are obtained by full fits of  $\mu_{\text{ex}}$  according to Eq. (12) and using Eqs. (13)–(15). Dotted lines are fits of  $\mu_{\text{ex}}$  under the constraint of constant  $\Delta C_p$ , that is,  $c = 0$ . Dashed lines are obtained by extrapolation, using Eqs. (1) and (2), from the reference temperature  $T_0 = 300$  K with  $\Delta S(T_0)$ ,  $\Delta H(T_0)$ , and  $\Delta C_p(T_0)$  taken from the full fit. The trajectories for the particle insertion are obtained from MD simulations of 895 SPC/E water molecules at constant pressure  $p = 1$  bar.

and a linearly varying heat capacity increment

$$\Delta C_p = d + 2cT, \quad (15)$$

see Fig. 4. To estimate how relevant the temperature dependence of the heat capacity is, we compare in Fig. 4 fit

results using the full expression (shown by solid lines) with constrained fits where the heat capacity is forced to be constant and thus  $c = 0$  (dotted lines in Fig. 4). We find the fits with  $c = 0$  to be of acceptable quality, which is in line with the fact that the actual heat capacity variation with temperature is small. This is quite important in light of the common assumption of a constant heat capacity used for extrapolating experimental solvation data. To check this procedure explicitly, we extrapolate  $\mu_{\text{ex}}(T)$ ,  $\Delta S(T)$ , and  $\Delta H(T)$ , using Eqs. (1) and (2) with a reference temperature  $T_0 = 300$  K, where we extract values for  $\Delta S(T_0)$ ,  $\Delta H(T_0)$ , and  $\Delta C_p(T_0)$  from the full fit. The resulting extrapolated curves (dashed lines) differ from the full fits for high temperatures, but otherwise the accuracy seems acceptable. The entropy shows convergence at a temperature around  $T_S^* \approx 350$  K, regardless of the method used to fit the data, which is gratifying as most experimental convergence temperatures are obtained via extrapolation. The value for  $T_S^*$  is significantly lower than the convergence temperature of about  $T_S^* \approx 400$  K found previously for hard spheres in SPC water.<sup>27</sup> We will discuss the origin of this discrepancy in detail further below but briefly mention here that it has to do with the different densities at which the simulations in Ref. 27 have been performed. The enthalpy convergence temperature lies outside the studied temperature range and occurs at roughly  $T_H^* \approx 205$  K. The difference in convergence temperature of about  $T_S^* - T_H^* \approx 140$  K is in rough agreement with experiments for simple liquid hydrocarbons, which give  $T_S^* \approx 380$  K and  $T_H^* \approx 290$  K, but there is a significant shift and we note that convergence in hydrocarbons occurs at vanishing entropy and enthalpy.<sup>10</sup> This is due to the fact that for homologous alkanes, convergence is mostly a consequence of additivity. For noble gases the entropy converges also around  $T_S^* \approx 380$  K, but the convergence entropy is nonzero. We will come back to this issue further below.

In Fig. 5 we plot the solvation entropy  $\Delta S(T)$  and the solvation enthalpy  $\Delta H(T)$  as a function of the heat capacity change  $\Delta C_p(T)$  for several different reference temperatures  $T = 300, 320, 340, 360$ , and  $380$  K. The convergence temperatures  $T_S^*$  and  $T_H^*$  and the entropies  $\Delta S^*$  and enthalpies  $\Delta H^*$  at convergence are summarized in Table I. For the entropy, the resultant numbers do not depend significantly on the reference temperature  $T$  and agree very well with what one directly reads off from Fig. 4. For the enthalpy, the individual linear fits in Fig. 5 are excellent, but the resulting convergence temperatures  $T_H^*$  and convergence enthalpies  $\Delta H(T)$  depend quite a bit on the reference temperature chosen. This is in line with the fact that the enthalpy data show less pronounced convergence when compared to the entropy, see Fig. 4. This shows clearly that a plot of  $\Delta S(T)$  or  $\Delta H(T)$  versus  $\Delta C_p(T)$  can be quite misleading since if the heat capacity change is not independent of the temperature the values for the convergence temperature and convergence thermodynamic quantity depend considerably on the reference temperature chosen.

In all what follows we use the full expression Eq. (12) with a temperature dependent heat capacity change to fit the particle insertion data. As a simple definition for the convergence temperature  $T_S^*$  we take the temperature at which the

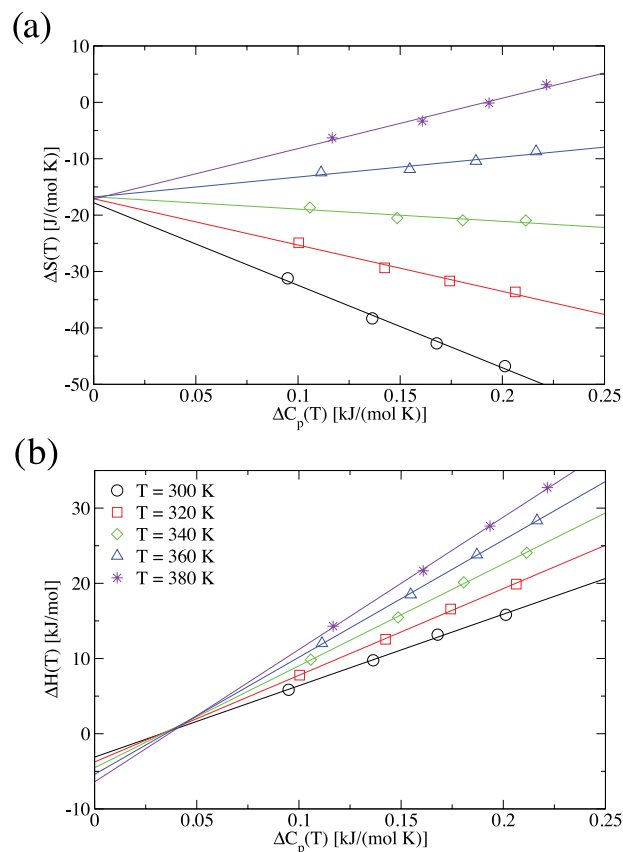


FIG. 5. (a) Solvation entropy  $\Delta S(T)$  and (b) solvation enthalpy  $\Delta H(T)$  of hard-core solutes of varying size as a function of the heat capacity change  $\Delta C_p(T)$  for different reference temperatures  $T = 300, 320, 340, 360$ , and  $380$  K. Data for entropies, enthalpies, and heat capacity changes are obtained from fits to the particle insertion data according to Eqs. (12)–(14). Solid lines show linear fits to the data according to Eqs. (1) and (2). The resulting convergence temperatures, entropies, and enthalpies are summarized in Table I. All data are obtained for SPC/E water at  $p = 1$  bar.

entropies of solutes with radii  $\sigma = 0.28$  and  $0.345$  nm intersect, that is,

$$\Delta S^{\sigma=0.28 \text{ nm}}(T_S^*) = \Delta S^{\sigma=0.345 \text{ nm}}(T_S^*). \quad (16)$$

The enthalpy convergence temperature  $T_H^*$  is defined analogously. We note that the convergence temperature depends systematically on the solute sizes that are compared,<sup>39</sup> so our definition only makes sense for the restricted range of radii actually considered by us.

## B. Influence of the stiffness of the solute-water repulsion

We first investigate the effect of the stiffness of the repulsive part of the potential, which is determined by the decay length  $1/B$ . For this purpose, we study purely repulsive solutes [i.e.,  $C = 0$  in Eq. (7)] with repulsion decay lengths in the range of  $1/B = 0.005$ – $0.03$  nm and compare with hard-core solutes. Figure 6(a) shows the excess chemical potential of a solute of radius  $\sigma = 0.33$  nm for different values of the decay length  $1/B$  (circles, full lines) and for a hard-core potential (squares, dashed line). This radius roughly corresponds to a methane molecule and is—because of the additivity of solvation properties—also relevant for alkane chains.

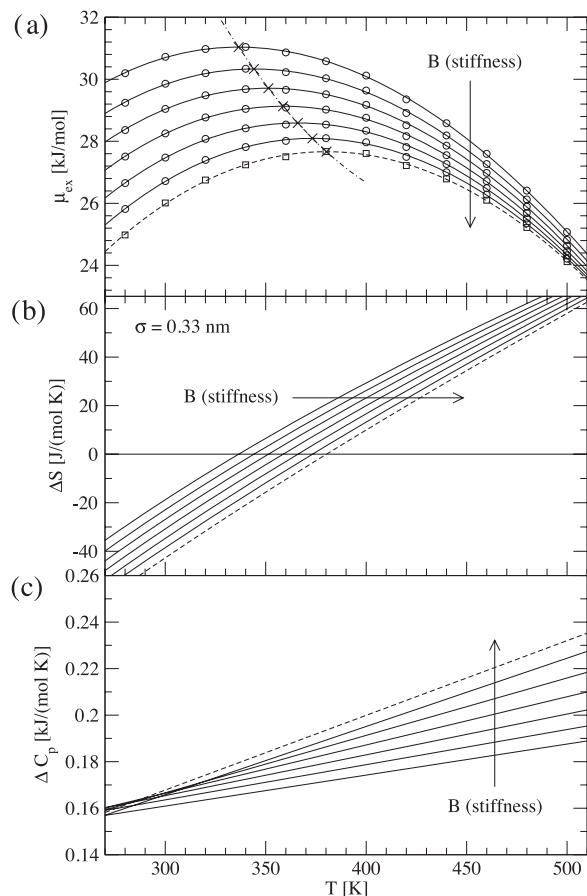


FIG. 6. (a) Excess chemical potential, (b) solvation entropy, and (c) heat capacity change upon solvation of purely repulsive spherical solutes of radius  $\sigma = 0.33$  nm for varying repulsion decay length  $1/B = 0.005, 0.01, 0.015, 0.02, 0.025$ , and  $0.03$  nm (circles, full lines) and a hard-core solute of the same radius (squares, dashed lines) as a function of the temperature. The symbols are results from particle insertion. Lines are fits to the data according to Eqs. (12)–(15). The crosses in (a) mark the positions of the maxima of the chemical potentials and the dashed-dotted line is a quadratic fit. The trajectories for the particle insertion are obtained by MD simulations of the SPC/E water model at a constant pressure of  $p = 1$  bar.

It is seen that the solvation free energy curves are systematically shifted downward for increasing potential stiffness, the stiffer the potential the less unfavorable the solvation process. This effect is more pronounced at small temperatures, leading to a shift of the free energy maxima to higher temperatures as indicated by the dashed-dotted line. The corresponding solvation entropies are shown in Fig. 6(b). For increasing potential stiffness they are shifted toward higher temperatures, while the temperature dependence of the heat capacity in Fig. 6(c) becomes more pronounced with increasing potential stiffness. One notes that there is no crossing of the curves, i.e., for solutes characterized by same radii but different stiffness one does not find entropy convergence. The entropy shifts have a pronounced influence on the convergence temperature, as shown in Fig. 7. For increasing stiffness the convergence is shifted to higher temperatures, reaching a value of  $T_S^* = 347$  K for the limiting case of a hard-core solute already shown in Fig. 4. Actual interaction potentials between solutes and water are far from the hard-core limit. For a realistic value of  $1/B = 0.03$  nm the convergence temper-

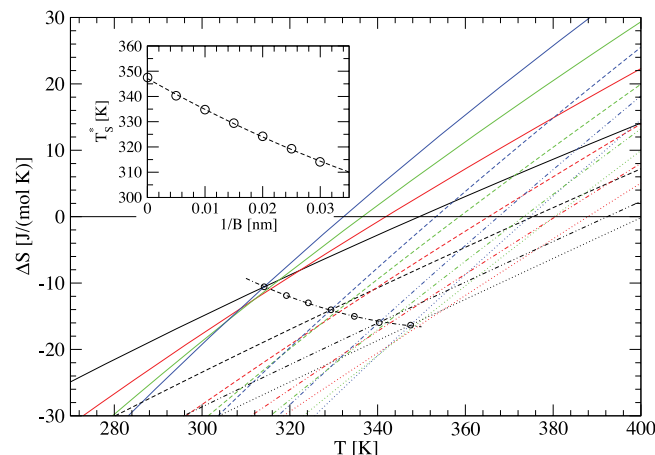


FIG. 7. Solvation entropies for purely repulsive solutes of varying size  $\sigma = 0.28$  nm (black),  $0.31$  nm (red),  $0.33$  nm (green), and  $0.345$  nm (blue), and varying potential stiffness  $1/B = 0.03$  nm (solid lines),  $0.015$  nm (dashed-dotted lines), and hard core (dotted lines) in SPC/E water at constant pressure  $p = 1$  bar. Circles indicate the intersection of the entropies of solutes with radii  $\sigma = 0.28$  and  $0.345$  nm for the same values of the potential stiffness as shown in Fig. 6 (not all entropy curves are shown). In the inset the convergence temperature defined by Eq. (16) is shown as a function of the repulsion decay length  $1/B$ .

ature is  $T_S^* = 314$  K. This temperature is even farther away from the experimentally observed value of  $T_S^* \approx 385$  K for noble gases and proteins.<sup>7</sup> In essence, we see that a certain softness of the repulsive part of the solute–water interaction is relevant as it significantly shifts the convergence temperature to lower temperatures, but this realistic feature does not improve the comparison with the experimental convergence temperature. We will later see that an attractive component to the interaction will bring the convergence temperature up again.

### C. Influence of water models, comparison with information theory

Here, we investigate the origin of the discrepancies of our particle insertion simulation results, obtained with SPC/E water so far, with previous theoretical results, based on information theory in conjunction with simulations for SPC water.<sup>27</sup> Figures 8(a) and 8(b) show the solvation free energies and entropies for hard-core solutes of varying size in SPC/E, SPC, and TIP5P water including the information theory results of Ref. 27. Our data for the three different water models indicated by solid lines are all obtained by particle insertion simulations at ambient pressure of  $p = 1$  bar. While the excess chemical potentials agree rather well for temperatures around 300 K, they differ significantly at higher temperatures, which leads to a considerable spread in the resulting entropy convergence temperatures by about 30 K in Fig. 8(b). The data from Ref. 27, shown by the orange dashed lines, were obtained by the information theory approach based on SPC simulations at densities corresponding to the experimental vapor-liquid coexistence curve of water. The deviations from all the other data at  $p = 1$  bar are significant. To understand the reason for this deviation, we resimulated SPC water at densities that match the experimental liquid–vapor coexistence



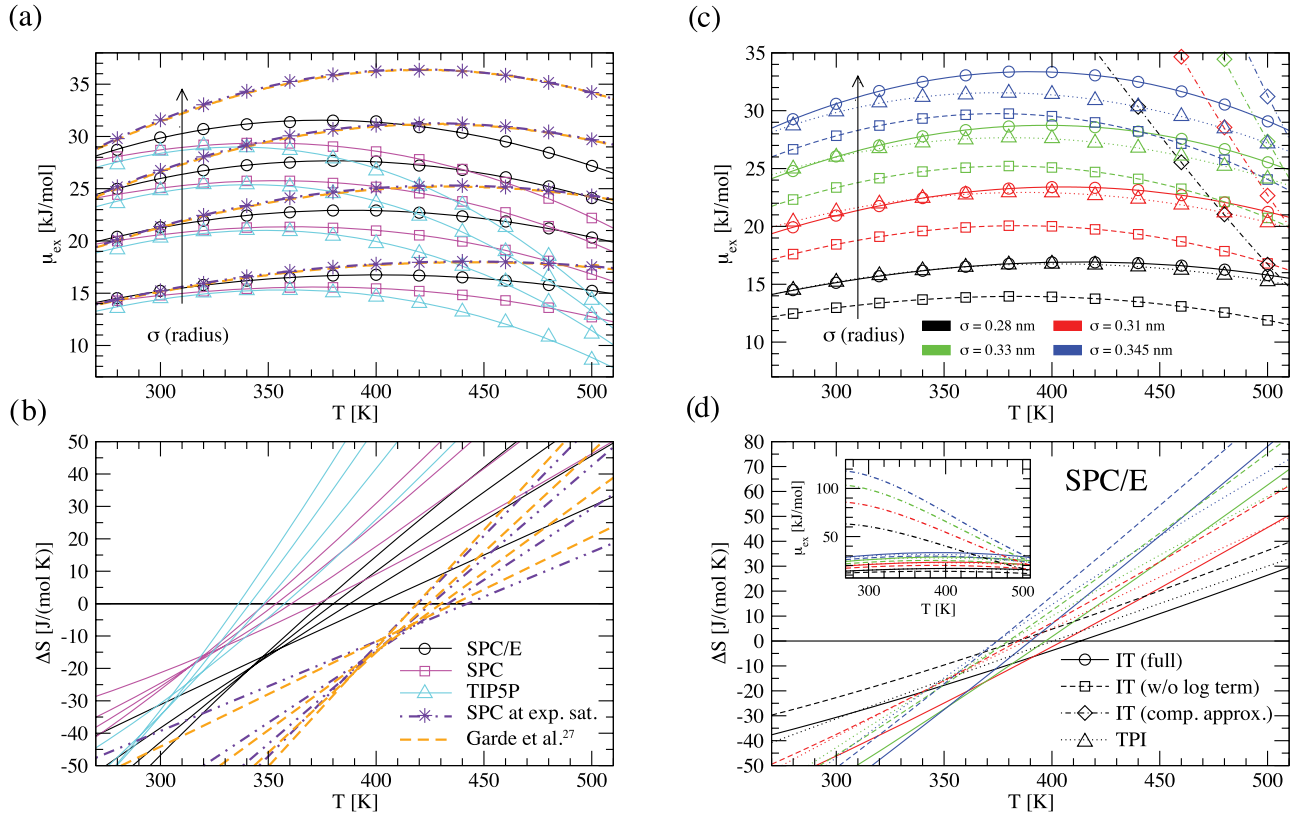


FIG. 8. (a) Excess chemical potentials  $\mu_{\text{ex}}$  and (b) solvation entropies of hard-core solutes of varying size  $\sigma = 0.28, 0.31, 0.33$ , and  $0.345$  nm for different water models. Symbols show the data obtained from particle insertion for the SPC/E (circles, black), the SPC (squares, magenta), and the TIP5P (triangles, cyan) water model, solid lines are obtained by fits to the particle insertion data according to Eq. (12) and (13). All simulations are performed at a constant pressure of  $p = 1$  bar. For comparison the information theory (IT) results of Garde *et al.* (Ref. 27) (orange dashed lines) and our own IT results for SPC water simulated at the densities of the experimental saturation curve (purple stars and double-dotted dashed lines) are also included. (c) Comparison of the chemical potential obtained by the information theory model [see Eq. (17)] including (circles and full lines) and neglecting (squares and dashed lines) the logarithmic term, the compressibility approximation [Eq. (20)] (diamonds and dashed-dotted lines), and the results from the particle insertion (triangles and dotted lines) for SPC/E water at a constant pressure of  $p = 1$  bar. Lines are fits to the data points. The inset in (d) shows the results for  $\mu_{\text{ex}}$  on a larger scale. (d) Comparison of the solvation entropies obtained by the temperature derivative of the fits to the chemical potentials shown in (c). Line styles have the same meaning as in (c).

densities and used the same information theory approach, the results are shown by purple stars and lines. The agreement with the previous IT approach is perfect, yet the disagreement with the other simulations still needs to be explained.

To that end, let us briefly reconsider the information theory model for the solvation of hard-core solutes.<sup>27,28,64</sup> It relates the excess chemical potential  $\mu_{\text{ex}}$  of a hard-core solute to the solvent particle number fluctuations inside a cavity of the size and shape of the solute, assuming Gaussian fluctuations. Using a continuous Gaussian approximation for the particle number distribution, the excess chemical potential follows as

$$\mu_{\text{ex}}^{\text{IT}} = k_{\text{B}} T \rho^2 v^2 / 2 \langle \delta n^2 \rangle + k_{\text{B}} T \ln(2\pi \langle \delta n^2 \rangle) / 2, \quad (17)$$

where  $\rho$  is the number density of the solvent,  $v$  is the volume of the cavity,  $\delta n = n - \langle n \rangle$ , and  $n$  is the number of solvent particles inside the randomly placed cavity. The second moment of the particle number fluctuations,  $\langle \delta n^2 \rangle$ , is related to a double integral over the oxygen–oxygen radial distribution function  $g_{\text{OO}}(r)$  by

$$\langle n(n-1) \rangle = \rho^2 \int_v d^3 r \int_v d^3 r' g_{\text{OO}}(|\vec{r} - \vec{r}'|), \quad (18)$$

which thus yields a straightforward way to evaluate the expression Eq. (17) based on simulation results for the distribution function  $g_{\text{OO}}(r)$ . In the thermodynamic limit, i.e., for infinite cavity volume  $v$ , the relation

$$\frac{\langle \delta n^2 \rangle}{\langle n \rangle} = \rho k_{\text{B}} T \kappa_T \quad (19)$$

connects particle number fluctuations with the isothermal compressibility  $\kappa_T = -(1/V)(\partial V / \partial p)_T$ . Neglecting the logarithmic term in Eq. (17) and using the compressibility, one obtains the drastically simplified expression

$$\mu_{\text{ex}}^{\text{IT}} \approx \frac{v}{2\kappa_T}. \quad (20)$$

Figures 8(c) and 8(d) show the excess chemical potentials and entropies as obtained from Eq. (17) including (circles, full lines) and neglecting (squares, dashed lines) the logarithmic term and also the predictions from Eq. (20) (diamonds, dashed-dotted lines) in comparison with the particle insertion results (triangles, dotted lines) for the SPC/E water model. Compressibilities for the SPC/E water model are obtained by a numerical derivative [see Eq. (11)] and are summarized in Table II. While Eq. (17) agrees well with the particle insertion results for small solutes and low temperatures, it differs

considerably for larger solutes and high temperatures. This indicates a breakdown of the Gaussian approximation so that non-Gaussian contributions to the particle number fluctuations become important.<sup>65</sup> The logarithmic term is not very important for the entropies, as has been noted before<sup>27</sup> The results from the simplified compressibility expression in Eq. (20) are completely off the particle insertion results (see inset of Fig. 8), showing that Eq. (19) is not valid on the nanoscale.

To explain the different results for the different water models and the discrepancy with the data of Ref. 27 we rearrange Eq. (17) without the logarithmic term in the following way,

$$\mu_{\text{ex}}^{IT} \approx \frac{k_B T}{2} \left( \frac{\langle n(n-1) \rangle}{\langle n \rangle^2} - 1 + \frac{1}{v\rho} \right)^{-1}, \quad (21)$$

where we used the relation  $\langle n \rangle = v\rho$ . We show  $\rho$  and  $\langle n(n-1) \rangle / \langle n \rangle^2$  for the different water models and pressures as obtained from simulations in Fig. 9. As noted before,<sup>27</sup>  $\langle n(n-1) \rangle / \langle n \rangle^2$  shows only a weak dependence on temperature or pressure and does not differ much between the different water models [see Fig. 9(b)]. The density  $\rho$  on the other hand shows a pronounced temperature dependence, differs appreciably between the water models and also shows deviations from the experimental data. Comparing with Eq. (21) and noting that the neglect of the logarithmic term in Eq. (17) is not very serious, as shown in Figs. 8(c) and 8(d), one concludes that it is mostly the variation in  $\rho$  that causes the differences in the solvation free energy,<sup>27,37,40</sup> as seen in Figs. 8(a) and 8(b). Note that the compressibility  $\kappa_T$  obtained for SPC/E, shown in the inset of Fig. 9(a), shows considerable dependence on temperature and agrees overall quite well with experimental data. But, as shown above, its temperature dependence cannot be used as a simple explanation of the entropy convergence phenomena since Eq. (19) is too inaccurate to describe solvation of small solutes.

In summary, two different factors contribute to the shift of the entropy convergence temperature between our simulated particle insertion data and the published information theory results: First, the temperature dependence of the experimental coexistence density is quite different from the simulated density at ambient pressure (and as a side remark, this deviation is more pronounced with SPC than with SPC/E), see Fig. 9(a), while the experimental densities at ambient and saturation pressures are very close. Second, there is a system-

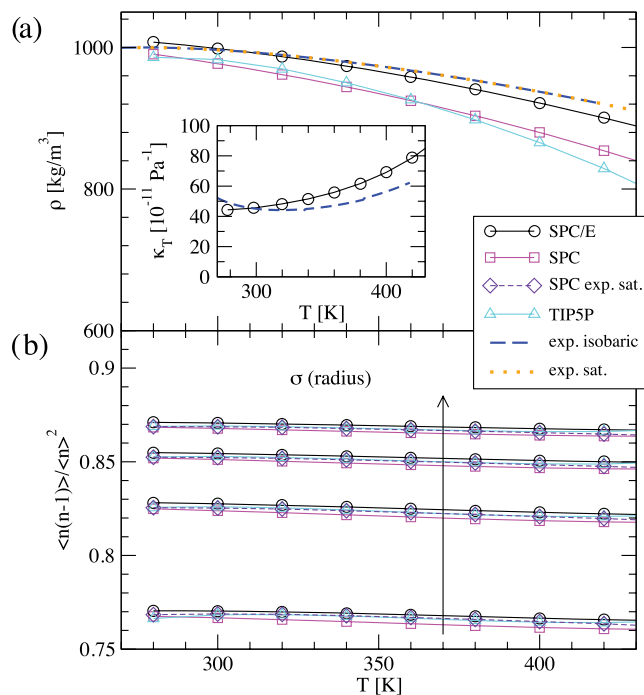


FIG. 9. (a) Water density  $\rho$  from MD simulations of the SPC/E (circles, black), the SPC (squares, magenta), and the TIP5P (triangles, cyan) water model at a pressure of  $p = 1$  bar. Also shown for comparison is the experimental water density for constant pressure  $p = 1$  bar (dashed line) (Ref. 62) and along the saturation curve (Ref. 63) (dotted line). The inset shows the compressibility of the SPC/E water model (circles) in comparison with the experimental data (dashed line) (Ref. 62). (b) Normalized solvent particle number fluctuations  $\langle n(n-1) \rangle / \langle n \rangle^2$  in a spherical cavity of varying radius  $\sigma = 0.28, 0.31, 0.33$ , and  $0.345$  nm for the three considered water models at constant pressure  $p = 1$  bar and for SPC water at the experimental saturation density (diamonds, purple). The symbols have the same meaning as in (a). Lines are guides to the eye.

atic shift of the entropy convergence to higher temperatures for the information theory results as compared to straight particle insertion results, see Figs. 8(c) and 8(d). Finally, we note that the coexistence densities of the SPC water model are significantly lower than the experimental ones,<sup>66</sup> simulating the SPC water model at experimental saturation densities therefore leads to unrealistically high pressures, e.g., 880 bar at  $T = 500$  K compared to the experimental saturation pressure of 26.4 bar.<sup>63</sup> It is not clear what the significance of such high pressure is, but it might significantly change the water properties and therefore lead to unrealistic results.

## D. Influence of attractive interactions

We next investigate the solvation of nonpolar solutes whose interaction with water also has an attractive part, according to Eq. (7). For a fixed decay length  $1/B = 0.02$  nm of the repulsive part of the potential, we vary the range of the attractive interaction ( $1/D = 0.05, 0.1$ , and  $0.2$  nm) as well as the potential depth ( $\epsilon_0 = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ , and  $0.8$  kJ/mol). In Figs. 10(a) and 10(b) we show the resulting solvation free energies and entropies for a solute of radius  $\sigma = 0.33$  nm and for  $1/B = 0.02$  nm and  $1/D = 0.05$  nm for different values of the potential depth  $\epsilon_0$ ; in Figs. 10(e) and 10(f) we show solvation free energies and

TABLE II. Isothermal compressibility  $\kappa_T$  of the SPC/E water model at a pressure of  $p = 1$  bar for varying temperature. The compressibilities are obtained from MD simulations by a numerical derivative according to Eq. (11). Error estimates are obtained by a block averaging algorithm.

T [K]	$\kappa_T$ [ $10^{-11}$ Pa $^{-1}$ ]
278	$44.32 \pm 0.17$
298	$45.50 \pm 0.17$
320	$48.23 \pm 0.19$
340	$51.53 \pm 0.20$
360	$55.91 \pm 0.25$
380	$61.61 \pm 0.28$
400	$69.10 \pm 0.44$
420	$78.66 \pm 0.96$

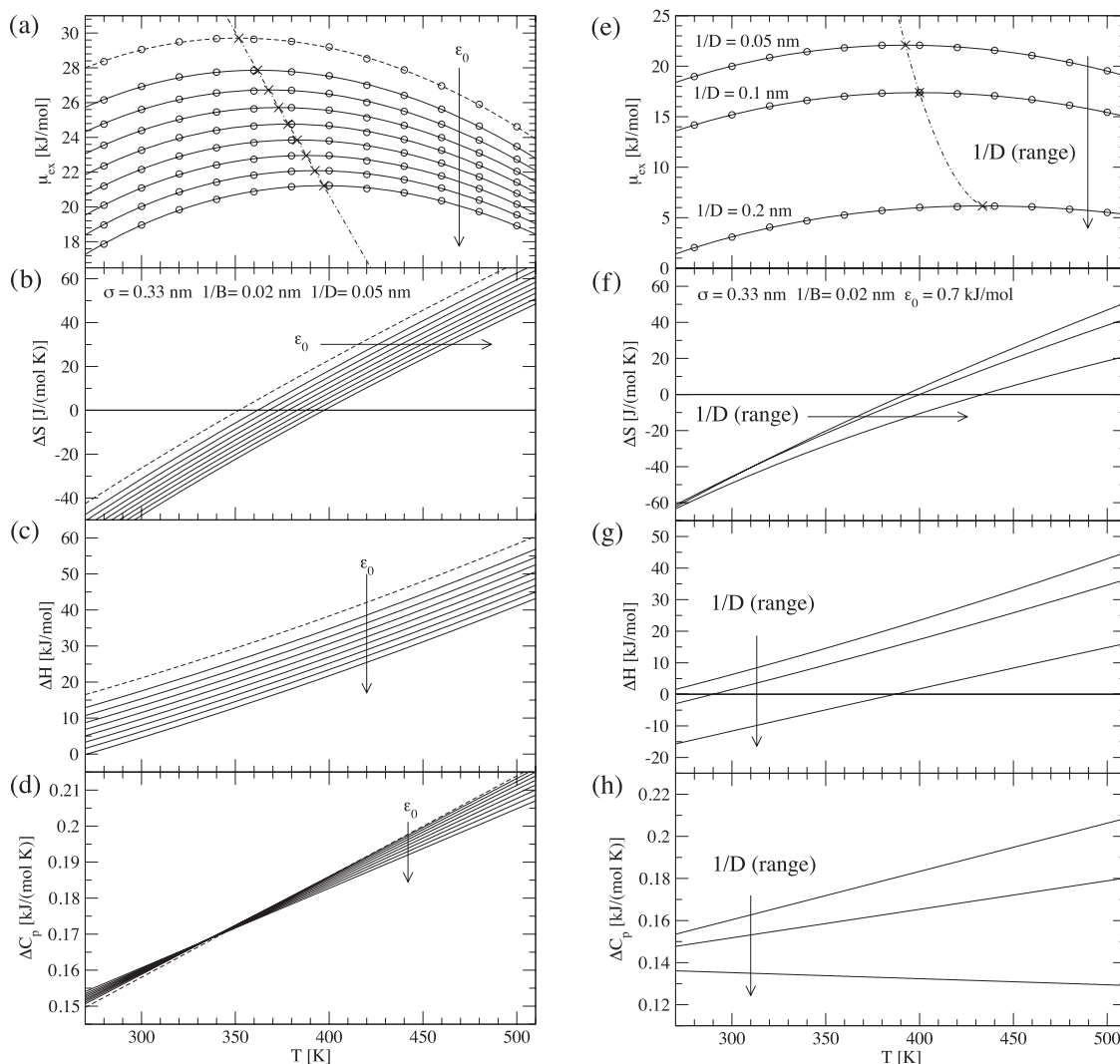


FIG. 10. Excess chemical potential for a spherical solute with radius  $\sigma = 0.33$  nm (a) for  $1/B = 0.02$  nm,  $1/D = 0.05$  nm and varying potential depth  $\epsilon_0 = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ , and  $0.8$  kJ/mol and (e) for  $1/B = 0.02$  nm,  $\epsilon_0 = 0.7$  kJ/mol and varying range of the attractive part of the potential  $1/D = 0.05, 0.1$ , and  $0.2$  nm. Circles show the data obtained by test particle insertion, lines are fits to the particle insertion data. The dashed line marks a purely repulsive solute ( $\epsilon_0 = 0$ ), while the solid lines correspond to the attractive solutes. Crosses indicate the position of the maxima of the free energy curves and the dashed-dotted line is a fit to the maxima. [(b), (f)] Solvation entropies, [(c), (g)] enthalpies and [(d), (h)] heat capacity changes obtained from derivatives of the fits to the free energies shown in (a) and (e). The trajectories for particle insertion are obtained by MD simulation of the SPC/E water model at  $p = 1$  bar.

entropies for  $1/B = 0.02$  nm and  $\epsilon_0 = 0.7$  kJ/mol for varying potential range  $1/D$ . It can be seen from Figs. 10(a) and 10(e), that for increasing potential depth and range the solvation free energies are shifted downward, as expected due to the more favorable enthalpic contribution of the attraction [see Figs. 10(c) and 10(g)]. For the largest range studied,  $1/D = 0.2$  nm, the solvation free energy approaches zero, meaning that we are almost describing hydrophilic (yet non-polar) solutes. At the same time the position of the maximum of the free energy curve is shifted toward higher temperatures. Accordingly, also the solvation entropy curves are shifted toward higher temperatures with increasing potential depth and range [see Figs. 10(b) and 10(f)]. Interestingly, for increasing attraction, that is as one goes from very hydrophobic solutes to less hydrophobic solutes, the solvation entropies become more negative while the heat capacities in Figs. 10(d) and 10(h) stay roughly the same. This stands in contrast to the common view according to which the solvation of hydrophilic

solutes is purely enthalpic and characterized by vanishing heat capacity.<sup>22,23</sup> We note, that the solvation free energies for a Lennard-Jones parameterization of methane<sup>40</sup> are well within the spread observed in Fig. 10(e). For example, at  $T = 300$  K Ashbaugh *et al.*<sup>40</sup> find a solvation free energy of  $9.3$  kJ/mol for OPLS-methane in SPC/E water, while for a methane size solute with  $\epsilon_0 = 0.7$  kJ/mol and  $1/B = 0.02$  nm we obtain  $15.2$  kJ/mol for  $1/D = 0.1$  nm and  $3.2$  kJ/mol for  $1/D = 0.2$  nm.

The effect of increasing potential depth and range on the convergence temperature is shown in Fig. 11, where we plot the solvation entropies and enthalpies for solutes of varying radii and different values of the potential depth  $\epsilon_0$  and range  $1/D$  as a function of the temperature. For the less stiff repulsive potential characterized by  $1/B = 0.02$  nm we show in Figs. 11(a) and 11(b) the intersection points of the entropy and enthalpy curves for solutes with radii  $\sigma = 0.28$  nm and  $\sigma = 0.345$  nm for attractive potential ranges of  $1/D = 0.05$  nm

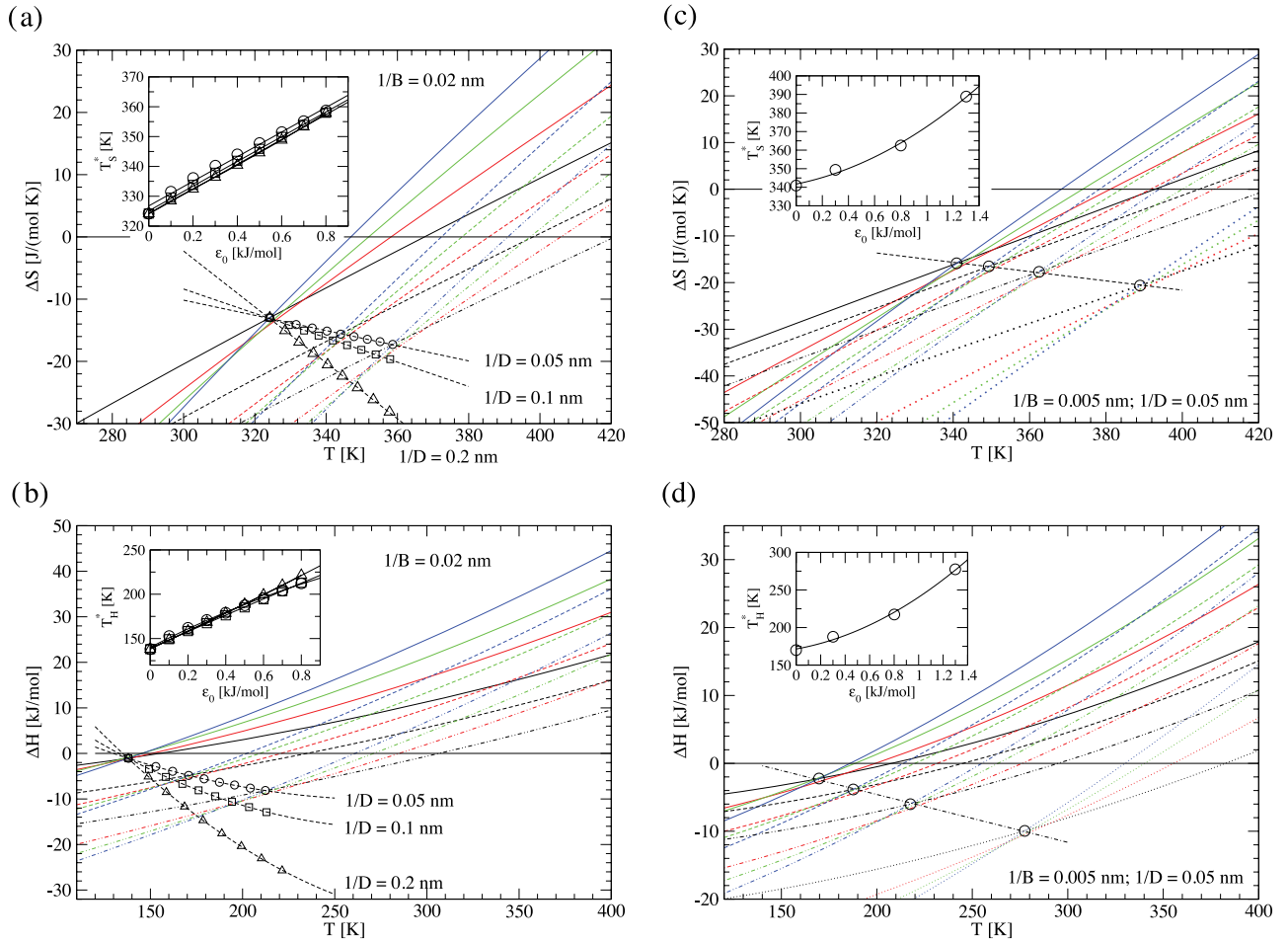


FIG. 11. (a) Solvation entropies for solutes with radii  $\sigma = 0.28$  nm (black),  $0.31$  nm (red),  $0.33$  nm (green), and  $0.345$  nm (blue) for  $1/B = 0.02$  nm,  $1/D = 0.05$  nm and varying potential depth  $\epsilon_0 = 0$  (solid lines),  $0.3$  kJ/mol (dashed lines), and  $0.8$  kJ/mol (dashed-dotted lines). The symbols show the intersection of the curves for solutes of size  $\sigma = 0.28$  nm and  $\sigma = 0.345$  nm for an attractive range  $1/D = 0.05$  nm (circles),  $1/D = 0.1$  nm (squares, entropy curves not shown), and  $1/D = 0.2$  nm (triangles, entropy curves not shown) and  $1/B = 0.02$  nm. The dashed curves are fits to the intersection points. The inset shows the convergence temperature defined by Eq. (16) for  $1/D = 0.05$  nm (circles),  $1/D = 0.1$  nm (squares) and  $1/D = 0.2$  nm (triangles) and  $1/B = 0.02$  nm. Solid curves show linear fits to the data. (b) Analogous results for the solvation enthalpy. [(c), (d)] Solvation entropy and enthalpy for the stiffer repulsive potential characterized by  $1/B = 0.005$  nm for varying potential depth  $\epsilon_0 = 0$  (solid lines),  $0.3$  kJ/mol (dashed lines),  $0.8$  kJ/mol (dashed-dotted lines) and  $1.3$  kJ/mol (dotted lines) and for only one attractive range  $1/D = 0.05$  nm.

(circles),  $1/D = 0.1$  nm (squares) and  $1/D = 0.2$  nm (triangles). For the latter two the entropy curves are not shown. In Figs. 11(c) and 11(d) we show data for  $1/D = 0.05$  nm and the stiffer repulsive potential  $1/B = 0.005$  nm. It can be seen that the convergence for increasing potential depth is shifted to higher temperatures, while the convergence entropies and enthalpies become more negative. This effect is more pronounced for more long-ranged attraction. In the insets of Fig. 11 we plot the convergence temperatures, defined by Eq. (16), as a function of the potential depth for varying potential ranges. Surprisingly, the convergence temperatures show very little dependence on the range of the attractive potential. It is apparent from Fig. 2(b), that a changing attractive potential depth also modifies the repulsive flank of the potential and thereby increases the apparent stiffness of the potential slightly. To understand the reason for the insensitivity of the entropy convergence temperature on the attractive range, we first have to disentangle the effects due to the repulsive and attractive parts of the potential. As a measure of the actual stiffness of the repulsive flank of the poten-

tial we define the effective stiffness  $1/B_{\text{eff}}$  of the potential [Eq. (7)] as

$$\frac{1}{B_{\text{eff}}} = -\frac{k_B T_0}{\partial U_{S-\text{OW}}(r)/\partial r|_{r=\sigma}}. \quad (22)$$

For a purely repulsive solute ( $C = 0$ ), this reduces to the usual decay length of the potential,  $1/B_{\text{eff}} = 1/B$ . The effective stiffness for attractive potentials with varying range  $1/D = 0.05, 0.1$ , and  $0.2$  nm is plotted in the inset of Fig. 12(b) as a function of the potential depth  $\epsilon_0$ . The increase in the effective stiffness with potential depth is smaller for the longer ranged potentials in accordance with Fig. 2. We assume, that the convergence temperature is an analytic function of the effective stiffness  $B_{\text{eff}}$ , the potential depth  $\epsilon_0$  and range  $D$ , that is  $T_S^* = T_S^*(B_{\text{eff}}, \epsilon_0, D)$ . Since with increasing the potential depth  $\epsilon_0$  also the effective stiffness  $B_{\text{eff}}$  increases, it is *a priori* not clear which of the two causes the change in the convergence temperature. In Fig. 12(a), we plot the convergence temperature for the attractive solutes as a function of  $1/B_{\text{eff}}$  and compare it to the results for the purely repulsive solutes.



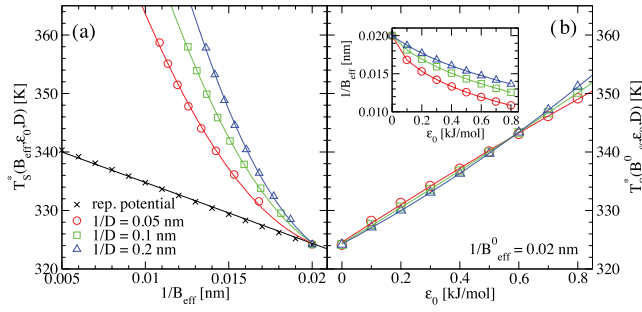


FIG. 12. (a) Convergence temperature for spherical solutes with  $1/B = 0.02$  nm,  $1/D = 0.05$  nm (circles),  $0.1$  nm (squares) and  $0.2$  nm (triangles) and varying depth ( $\epsilon_0 = 0.0 - 0.8$  kJ/mol) as a function of the effective stiffness  $(1/B)_{\text{eff}}$ , defined in Eq. 22. Lines are quadratic fits to the data. Also shown for reference is the convergence temperature for purely repulsive solutes (crosses) for varying  $B$  including a linear fit to the data. (b) Convergence temperature  $T_S^*$ , corrected for the varying effective stiffness as a function of the potential depth  $\epsilon_0$ . The inset shows the effective stiffness of the attractive potentials as a function of the potential depth. Symbols have the same meaning as in (a). All data are obtained by particle insertion in SPC/E water at  $p = 1$  bar.

It is seen that the convergence temperature for the attractive solutes rises much faster with increasing effective stiffness than for the purely repulsive solutes, so the change in effective stiffness alone cannot explain the shift in the convergence temperature; it is rather a combined effect of both factors. In order to disentangle the effects of the attractive interactions and the effective stiffness of the repulsive part of the potential, we expand  $T_S^*(B_{\text{eff}}, \epsilon_0, D)$  for fixed  $D$  in a Taylor series with respect to  $\epsilon_0$  and  $B_{\text{eff}}$ ,

$$T_S^*(B_{\text{eff}}, \epsilon_0, D) \approx T_S^*(B_{\text{eff}}^0, \epsilon_0^0, D) + \frac{\partial T_S^*(B_{\text{eff}}^0, \epsilon_0^0, D)}{\partial B_{\text{eff}}} (B_{\text{eff}} - B_{\text{eff}}^0) + \frac{\partial T_S^*(B_{\text{eff}}^0, \epsilon_0^0, D)}{\partial \epsilon_0} (\epsilon_0 - \epsilon_0^0). \quad (23)$$

Adding and subtracting  $T_S^*(B_{\text{eff}}^0, \epsilon_0^0, D)$  one can rearrange Eq. (23) to yield

$$T_S^*(B_{\text{eff}}, \epsilon_0, D) \approx T_S^*(B_{\text{eff}}^0, \epsilon_0^0, D) + T_S^*(B_{\text{eff}}^0, \epsilon_0, D) - T_S^*(B_{\text{eff}}^0, \epsilon_0^0, D). \quad (24)$$

If we choose  $1/B_{\text{eff}}^0 = 0.02$  nm and  $\epsilon_0^0 = 0$ , then  $T_S^*(B_{\text{eff}}^0, \epsilon_0^0, D)$  and  $T_S^*(B_{\text{eff}}^0, \epsilon_0, D)$  are given by the results for purely repulsive solutes and we can calculate  $T_S^*(B_{\text{eff}}^0, \epsilon_0, D)$  by solving Eq. (24). We plot  $T_S^*(B_{\text{eff}}^0, \epsilon_0, D)$  for  $1/B_{\text{eff}}^0 = 0.02$  nm in Fig. 12(b) as a function of the potential depth  $\epsilon_0$  for the three different ranges studied. Still a strong dependence on  $\epsilon_0$  can be observed, while there is almost no change with the range of the attractive part of the potential. We conclude that the entropy convergence temperature shows no dependence on the range of the attractive part of the potential and only depends on the attractive strength and repulsive stiffness. This suggests that the main contribution to the temperature dependence of the solvation free energy comes from the first solvation shell around the solute and is therefore determined by the short range structure of the hydrogen bonding network.

## V. CONCLUSION

We performed an analysis of the solvation properties of small spherical solutes for a broad class of solute-water potentials. In all cases the free energies show a maximum at a temperature that depends quite sensitively on details of the potential, below that temperature the solvation entropy is negative (i.e., the solute orders the water), above that temperature the entropy is positive (i.e., the solute increases the water disorder). The heat capacity is invariably positive. These are the common signatures of hydrophobic solvation of small solutes. No convergence is found for solutes of the same size characterized by different  $1/B$  and  $\epsilon_0$ , but restricted convergence is observed for the same  $1/B$  and  $\epsilon_0$  and different radii of the solutes. The entropy convergence temperature, at which the solvation entropies of solutes of different radii are the same, depends also on potential details in the following fashion: the less stiff the short-ranged repulsive part of the potential, the lower the convergence temperature, the deeper the attractive part, the higher the convergence temperature. The range of the attractive part does not influence the convergence properties, which suggests that it is only the first water shell that plays a role for convergence. Nevertheless, these results mean for the case of protein denaturation, where different amino acids plausibly are characterized by different effective water-amino acid potentials, entropy or enthalpy convergence cannot be expected in the general case. This is in agreement with a recent analysis of a large unbiased set of protein data.<sup>41</sup> It has been suggested that the spread in the protein convergence properties in Fig. 1 can be related to the size of the proteins.<sup>30</sup> However, when we plot the deviation from a straight-line fit as a function of the number of residues [see Fig. 1(b)] we seem not to obtain any systematic trend. Likewise, although we show that the extrapolation procedure which is used to obtain thermodynamic data at one temperature, which basically rests on the assumption of a constant heat capacity, is fine for the solvation of nonpolar solutes (see our analysis in Figs. 4 and 5), there could be extrapolation artifacts in the protein denaturation data. However, when we separate data from proteins that denature around  $60^\circ$  C (shown by open circles) from proteins that denature further away from  $60^\circ$  C (filled circles) we do not see a systematic difference in the two data sets, ruling out problems with the extrapolation method as a possible cause of the absence of entropy convergence. Taken together all evidence, entropy convergence of protein denaturation does not seem to be evidenced by the experimental data.

Although strictly speaking we cannot conclude that the breakdown of convergence in protein denaturation data is caused by the variations of the interactions of individual amino acids with water as we have studied in this paper, we note that these variations are sufficient to give rise to a breakdown of convergence. Other factors that would work in the same direction are specificities in the folded state, which we did not address.

Most directly our results are applicable to the solvation of gaseous particles in water, since here specifics of the reference state (which is the liquid state for liquid hydrocarbons or the protein core for folded proteins) are absent. We

stress again a subtle point mentioned already in the previous sections: for liquid hydrocarbons which form a homologous series, entropy convergence is foremost a consequence of the additivity of solvation properties of individual methyl and methylene groups and yields a vanishing entropy at the convergence temperature. In contrast, convergence for different noble gases is not connected to the additivity hypothesis but rather to an insensitivity of solvation thermodynamics to variations in the solute–water interaction and leads in general to a nonvanishing entropy at convergence.

We see that the entropy convergence of noble gases, which experimentally occurs around 380 K, is quite well reproduced for sensible values of the attractive strength and range around  $1/B = 0.005$  nm and  $\epsilon_0 = 1.3$  kJ/mol. Experimentally, the convergence is far from perfect, since, e.g., the entropies of Ne, Ar, and Xe do not cross at one temperature. The actual value of the entropy at convergence comes also out quite well from our simulation results. As seen in Fig. 2 of Ref. 39, Ar and Xe cross at an entropy of  $\Delta S \approx -4$  cal/mol K  $\approx -17$  J/mol K, which is in line with typical convergence entropies we obtain for a wide range of values of the attractive depth between  $\epsilon_0 = 0.3$  and 1.3 kJ/mol, see Fig. 11(c).

Note that the model potential considered by us should be viewed as realistic for the solvation of hydrophobic particles which cannot form hydrogen bonds with the water. In the future, it will be interesting to study polar solutes as well and to check whether robust and significant differences to the solvation of unpolars solutes are obtained.

## ACKNOWLEDGMENTS

Financial support by the Elitenetzwerk Bayern in the framework of the doctorate program Materials Science of Complex Interfaces and the Deutsche Forschungsgemeinschaft (DFG Grant No. NE 810/7), and the Ministry for Economy and Technology (BMW) in the framework of the AiF project “Simulation and prediction of salt influence on biological systems” is gratefully acknowledged. The authors thank the Leibniz-Rechenzentrum München (LRZ) for computing time on HLRB-II (Project pr28xe).

- <sup>1</sup>J. Buchner and T. Kiefhaber, *Protein Folding Handbook*, 1st ed. (Wiley-VCH, Weinheim, 2005).
- <sup>2</sup>C. Anfinsen, *Science* **181**, 223 (1973).
- <sup>3</sup>W. Kauzmann, *Nature* **325**, 763 (1987).
- <sup>4</sup>K. A. Dill, *Biochemistry* **29**, 7133 (1990).
- <sup>5</sup>W. Kauzmann, *Adv. Protein Chem.* **14**, 1 (1959).
- <sup>6</sup>P. L. Privalov and S. J. Gill, *Adv. Protein Chem.* **39**, 191 (1988).
- <sup>7</sup>K. P. Murphy, P. L. Privalov, and S. J. Gill, *Science* **247**, 559 (1990).
- <sup>8</sup>P. L. Privalov and N. N. Khechinashvili, *J. Mol. Biol.* **86**, 665 (1974).
- <sup>9</sup>P. L. Privalov, *Adv. Protein Chem.* **33**, 167 (1979).
- <sup>10</sup>R. L. Baldwin, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8069 (1986).
- <sup>11</sup>S. J. Gill, N. Nichols, and I. Wadsö, *J. Chem. Thermodyn.* **8**, 445 (1976).
- <sup>12</sup>E. Wilhelm, R. Battino, and R. Wilcock, *Chem. Rev.* **77**, 219 (1977).
- <sup>13</sup>S. Dec and S. J. Gill, *J. Solution Chem.* **13**, 27 (1984).
- <sup>14</sup>H. Naghibi, S. Dec, and S. J. Gill, *J. Phys. Chem.* **90**, 4621 (1986).
- <sup>15</sup>H. Naghibi, S. Dec, and S. J. Gill, *J. Phys. Chem.* **91**, 245 (1987).
- <sup>16</sup>H. Naghibi, D. Ownby, and S. J. Gill, *J. Chem. Eng. Data* **32**, 422 (1987).
- <sup>17</sup>S. Dec and S. J. Gill, *J. Solution Chem.* **14**, 417 (1985).
- <sup>18</sup>G. Olofson, A. Oshodj, E. Qvarnström, and I. Wadsö, *J. Chem. Thermodyn.* **16**, 1041 (1984).
- <sup>19</sup>K. P. Murphy and S. J. Gill, *Thermochim. Acta* **172**, 11 (1990).
- <sup>20</sup>K. P. Murphy and S. J. Gill, *J. Chem. Thermodyn.* **21**, 903 (1989).

- <sup>21</sup>K. P. Murphy, *Biophys. Chem.* **51**, 311 (1994).
- <sup>22</sup>B. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5154 (1991).
- <sup>23</sup>K. P. Murphy and S. J. Gill, *J. Mol. Biol.* **222**, 699 (1991).
- <sup>24</sup>A. Yang, K. A. Sharp, and B. Honig, *J. Mol. Biol.* **227**, 889 (1992).
- <sup>25</sup>L. Fu and E. Freire, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9335 (1992).
- <sup>26</sup>B. Guillot and Y. Guissani, *J. Chem. Phys.* **99**, 8075 (1993).
- <sup>27</sup>S. Garde, G. Hummer, A. E. Garcia, M. E. Paulaitis, and L. R. Pratt, *Phys. Rev. Lett.* **77**, 4966 (1996).
- <sup>28</sup>G. Hummer, S. Garde, A. E. Garcia, M. E. Paulaitis, and L. R. Pratt, *J. Phys. Chem. B* **102**, 10469 (1998).
- <sup>29</sup>S. Garde, A. E. Garcia, L. R. Pratt, and G. Hummer, *Biophys. Chem.* **78**, 21 (1999).
- <sup>30</sup>D. M. Huang and D. Chandler, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8324 (2000).
- <sup>31</sup>G. Hummer, S. Garde, A. E. Garcia, and L. R. Pratt, *Chem. Phys.* **258**, 349 (2000).
- <sup>32</sup>S. Garde and H. S. Ashbaugh, *J. Chem. Phys.* **115**, 977 (2001).
- <sup>33</sup>H. S. Ashbaugh, T. M. Truskett, and P. G. Debenedetti, *J. Chem. Phys.* **116**, 2907 (2002).
- <sup>34</sup>D. M. Huang and D. Chandler, *J. Phys. Chem. B* **106**, 2047 (2002).
- <sup>35</sup>G. Graziano and B. Lee, *Biophys. Chem.* **105**, 241 (2003).
- <sup>36</sup>B. Widom, P. Bhimalapuram, and K. Koga, *Phys. Chem. Chem. Phys.* **5**, 3085 (2003).
- <sup>37</sup>D. Paschek, *J. Chem. Phys.* **120**, 6674 (2004).
- <sup>38</sup>G. Graziano, *J. Phys. Chem. B* **109**, 12160 (2005).
- <sup>39</sup>H. S. Ashbaugh and L. R. Pratt, *Rev. Mod. Phys.* **78**, 159 (2006).
- <sup>40</sup>H. S. Ashbaugh, N. J. Collett, H. W. Hatch, and J. A. Staton, *J. Chem. Phys.* **132**, 124504 (2010).
- <sup>41</sup>A. D. Robertson and K. P. Murphy, *Chem. Rev.* **97**, 1251 (1997).
- <sup>42</sup>D. M. Huang, P. L. Geissler, and D. Chandler, *J. Phys. Chem. B* **105**, 6704 (2001).
- <sup>43</sup>W. Becktel and J. A. Schellman, *Biopolymers* **26**, 1859 (1987).
- <sup>44</sup>J. Sturtevant, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2236 (1977).
- <sup>45</sup>K. P. Murphy, V. Bhakuni, D. Xie, and E. Freire, *J. Mol. Biol.* **227**, 293 (1992).
- <sup>46</sup>C. Chothia, *J. Mol. Biol.* **105**, 1 (1976).
- <sup>47</sup>S. J. Gill and I. Wadsö, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2955 (1976).
- <sup>48</sup>C. Y. Hu, H. Kokubo, G. C. Lynch, D. W. Bolen, and B. M. Pettitt, *Protein Sci.* **19**, 1011 (2010).
- <sup>49</sup>H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
- <sup>50</sup>H. J. C. Berendsen, J. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces: Proceedings of the 14th Jerusalem Symposium on Quantum Chemistry and Biochemistry*, edited by B. Pullman (Reidel, Dordrecht, The Netherlands, 1981), p. 331.
- <sup>51</sup>M. Mahoney and W. L. Jorgensen, *J. Chem. Phys.* **112**, 8910 (2000).
- <sup>52</sup>J. R. Errington and A. Z. Panagiotopoulos, *J. Phys. Chem. B* **102**, 7470 (1998).
- <sup>53</sup>B. Widom, *J. Chem. Phys.* **39**, 2808 (1963).
- <sup>54</sup>P. Sindzingre, G. Ciccotti, C. Massobrio, and D. Frenkel, *Chem. Phys. Lett.* **136**, 35 (1987).
- <sup>55</sup>K. A. Motakabbir and M. L. Berkowitz, *J. Phys. Chem.* **94**, 8359 (1990).
- <sup>56</sup>B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *J. Chem. Theory Comput.* **4**, 435 (2008).
- <sup>57</sup>D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, *J. Comput. Chem.* **26**, 1701 (2005).
- <sup>58</sup>H. J. C. Berendsen, J. Postma, W. F. van Gunsteren, A. Dinola, and J. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- <sup>59</sup>U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995).
- <sup>60</sup>T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- <sup>61</sup>M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 1st ed. (Oxford University Press, Oxford, 1987).
- <sup>62</sup>G. S. Kell, *J. Chem. Eng. Data* **20**, 97 (1975).
- <sup>63</sup>E. Lemmon, M. McLinden, and D. Friend, “Thermophysical Properties of Fluid System,” in *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, edited by P. J. Linstrom and W. G. Mallard (National Institute of Standards and Technology, Gaithersburg, MD); see <http://webbook.nist.gov>, 2010.
- <sup>64</sup>G. Hummer, S. Garde, A. E. Garcia, A. Pohorille, and L. R. Pratt, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8951 (1996).
- <sup>65</sup>A. J. Patel, P. Varilly, and D. Chandler, *J. Phys. Chem. B* **114**, 1632 (2010).
- <sup>66</sup>J. de Pablo, J. Prausnitz, H. Strauch, and P. Cummings, *J. Chem. Phys.* **93**, 7355 (1990).